

...these notes may be updated/improved/ as the lecture proceeds: make sure you always use the latest version, as may be seen from the date on the cover...

Space astrophysics: Error analysis and fitting

Frank Verbunt

May 10, 2017

In this lecture we discuss how one can describe observations, and how one can compare measurements with model predictions. We start with some preparatory discussions about commonly occurring distributions, and discuss various definitions of the *average* and the spread around the average. We also discuss error propagation.

Next, we discuss the criteria that should be used to decide which fit of a data set is considered to be the best, for the case where the errors are distributed according to the Gauss-function. We describe fitting procedures specific for the least squares criterion, both for the linear case and for the non-linear case.

We also discuss best-fit criteria for the case where the errors are Poissonian, and two fitting procedures which are general, and which may be applied both for least squares and for the case where the errors follow a Poissonian distribution.

The lecture notes also contain a practical part, which explains simple programming in *Python*

INTERNATIONAL

Concorde's Fall in Safety Ratings Shows Data Can Be

By Don Phillips
Washington Post Service

WASHINGTON — The crash of the Air France Concorde last week transformed the safest aircraft from the safest plane on Earth to the most dangerous, statistically speaking.

This statistical quirk shows the dangers of measuring aviation safety by numbers alone in an era when flying is so safe that just one accident can skew the ratings dramatically. But statistics can also tell useful stories.

For more than three decades, the Concorde flew with no crashes, no deaths and no injuries more serious than bumps and bruises from occasional evacuations after nonfatal incidents. That means that on the industry's standard safety measure, "hull losses" per million flights, it scored a perfect zero. A hull loss is counted when an

airplane is damaged so badly that it will never fly again.

And because the Concorde has been in service far longer than other aircraft that now have zero hull-loss ratings — the Airbus A330 and A340 and the Boeing 777, 737NG and 717 — many people considered its record the best.

But because there are so few Concordes and because each flies fewer than 1,000 hours a year, the crash Tuesday raised its hull loss per million flights figure to 11.64, according to statistics issued by Boeing Co.

This is by far the worst record among jetliners flying today. It is exceeded only by those of the first generation of jets, which have long since been phased out — the Comet, the Caravelle, the Trident and the VC-10. Together they racked up an average figure of 15.51.

"I am bothered by continual references in all the

media about the safety record of the Concorde," said John Purvis, Boeing's former safety chief who is now a consultant and partner with Safety Services International of Seattle. "The entire fleet was good up to now, but in fact, it had less than 80,000 takeoffs and landings."

"I think people have to realize that safety is a numbers game," Mr. Purvis said. "Concorde had a good record partly because there weren't many of them, they were underutilized and had special maintenance handling and crews that were the cream of the crop."

The Boeing 737, by contrast, has had 77 crashes but still has an excellent safety record, ranging from 1.25 per million to 0.43 to zero hull losses for different versions of the plane, because it is the world's most common airliner. The huge 737 fleet flies more hours in one week than the Concordes have flown in their entire existence.

The 737 has great exposure to potential crash situations but still avoids accidents.

At first glance, the wide-body McDonnell Douglas MD-11, with five crashes and a hull-loss rate of 6.54 per million departures, appears to be the least safe subsonic aircraft now flying. This is worse than the first-generation U.S. airliners — the Boeing 707 has 116 hull losses and a rate of 6.51, while the McDonnell Douglas DC-8 has 72 losses and a rate of 5.91.

But a more careful examination tells a different story. First, the MD-11 has not yet flown a million departures. This makes it also susceptible to statistical quirks. And several of the crashed MD-11s were cargo planes, meaning there was no danger to the flying public.

The 707 and DC-8, on the other hand, suffered from the growing pains of the early jet age. Now, most of them are flying in the Third World, where

Misleading

crash rates tend to be higher for all types of planes.

■ Fault Delays British Concorde Flight

A British Airways Concorde flight to New York was delayed for more than an hour at London Heathrow Airport on Sunday after a technical fault forced the airline to use a backup plane, The Associated Press reported.

After a fault was discovered with a refueling system on the aircraft that was scheduled to make the flight, another Concorde jet was used instead, an airline representative said. "The passengers — there were 51 of them — hadn't actually boarded, so we put them on a substitute aircraft," she said. "We always have a standby Concorde ready."

She said the technical fault was specific to the plane involved and did not mean the rest of the fleet had to be checked.

Figure 1: How not to do statistics...

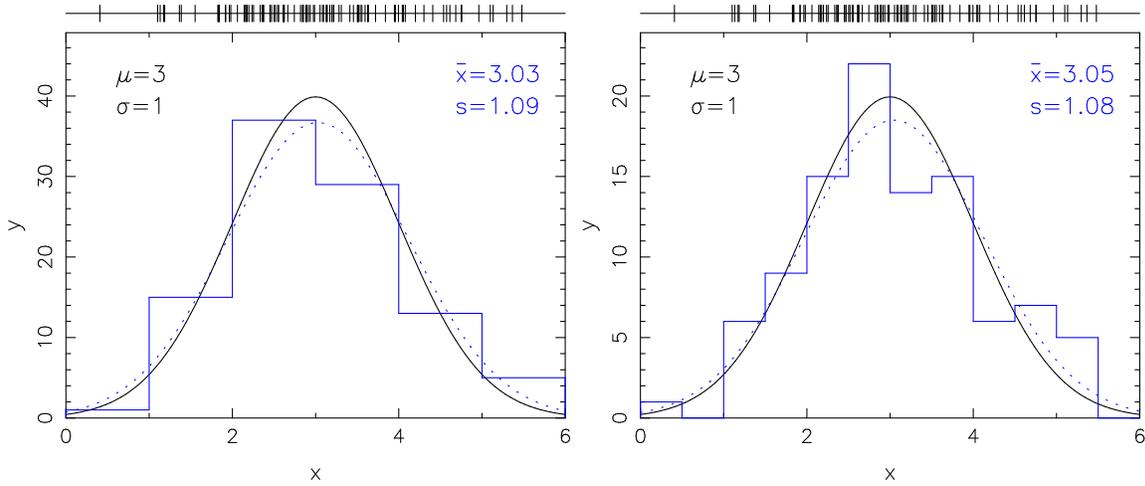


Figure 2: Illustration of the difference between the parent distribution (solid line) and the distribution (dotted line) derived from a finite number (in this case 100) of measured values (histogram). The parent distribution is a Gaussian (Eq. 1) with average $\mu = 3.0$ and width $\sigma = 1.0$. The parameters derived with Eqs. 7, 8 from the unbinned data (shown as vertical lines on top) are $\bar{x} = 3.03$ and $s = 1.06$. The parameters derived after binning depend on the number of bins (!), even though the measurement set is identical

1 Parent distribution: the concept

We may discriminate between the correctness and the precision of an experiment. The *correctness* of an experiment is a measure of how close the result of an experiment comes to the true value. Therefore, it is a measure of the correctness of the result. The *precision* of an experiment is a measure of how exactly the result is determined, without reference to what that result means. It is also a measure of how reproducible the result is. The *absolute precision* indicates the magnitude of the uncertainty in the result in the same units as the result. The *relative precision* indicates the uncertainty in terms of a fraction of the value of the result.

Suppose we have a quantity x which we measure. If we measure it many times, the measurements will show a certain distribution. The distribution that one would obtain by measuring an infinite number of times is called the parent distribution. The more often one measures, the better this distribution is approximated. In practice, of course, one measures a finite number of times, and thus quantities of the parent distribution can only be determined approximately.

It is general usage to indicate quantities of the parent distribution with Greek letters (e.g. μ, σ) and quantities determined from a finite number of measurements with Latin letters (e.g. \bar{x}, s).

As an illustration consider a Gaussian distribution with an average μ and width σ :

$$G(x, \mu, \sigma) \equiv \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (1)$$

This distribution is shown in Figure 2, for $\mu = 3$ and $\sigma = 1$. We now simulate a measurement of the quantity x by drawing 100 random values from this distribution, and bin them in bins of width 0.5. The distribution of the measurements is also shown in the figure. We estimate – with Eqs. 7, 8 below – the average and width of the

parent function, and find that the values of \bar{x} as estimate for μ and s as estimate for σ , depend on whether we bin, and on the number of bins. The Gaussian distributions with these values for μ and σ are also shown in the figure, and differ somewhat from the actual distribution.

It is useful to discriminate two reasons why we obtain a *distribution* of the measurement results, rather than one single value. One reason may be that the quantity we measure is intrinsically distributed. If we determine the income of arbitrarily selected Dutch people, we will obtain a range of values because there is a range of incomes, i.e. not all incomes are the same. Another reason may be that the measurement has a finite accuracy. If we determine the charge of an electron, we will find a distribution of values which reflects the precision of the measurement. In many measurements, both reasons apply. Whichever is the dominant reason, we can always describe the set of measurements in terms of an average and a spread around the average, and it is with this that we will start.

1.1 Characterizing a distribution with average and spread

Suppose we have a distribution of the measurable quantity x . As illustrated in Figure 2, a finite number of measurements will only approximate the parent distribution. In general we want to know what probable or characteristic values of x are, and what the spread of individual measurements x_i is.

We first define several characteristic values for the distribution.

The average or mean μ of the parent distribution is given by

$$\mu \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i \quad (2)$$

The median value $\mu_{1/2}$ of the parent distribution is that value at which there are equal numbers of values higher and lower than it. Expressed differently, the probability that an arbitrarily selected value is higher than the median is 50%, and the probability that an arbitrarily selected value is lower is 50% also:

$$P(x_i > \mu_{1/2}) \equiv P(x_i < \mu_{1/2}) \equiv 0.5 \quad (3)$$

Finally, the most probable value μ_{max} of the parent distribution is the value at which the distribution peaks:

$$P(\mu_{max}) \geq P(x_i \neq \mu_{max}) \quad (4)$$

For symmetric distributions with a single maximum, like the Gauss function, the average, median and most probable value are all the same. They are generally different in asymmetric functions, as for example in the Maxwellian (Figure 3).

To know how broad a distribution is, we can also define various quantities that indicate the width. A simple one is the average absolute value of the deviation from the average, i.e.

$$\alpha \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N |x_i - \mu| \quad (5)$$

The absolute value is required, because the straightforward average deviation from the average is zero by definition. Whereas α is a useful quantity, it is difficult to handle mathematically because of the occurrence of an absolute value.

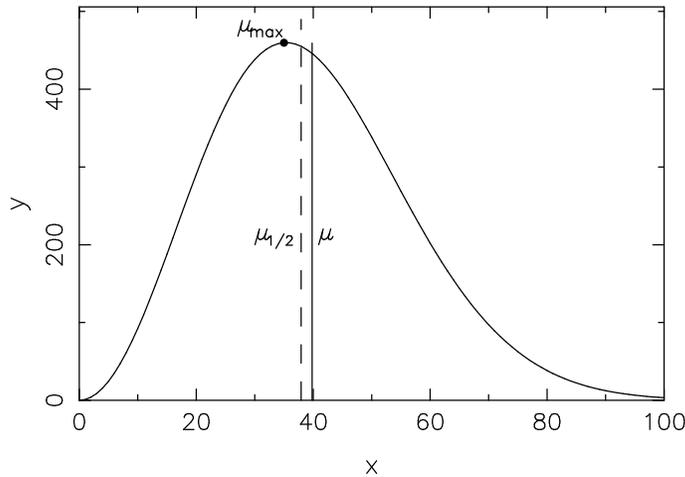


Figure 3: *Illustration of various definitions of the average, illustrated for the Maxwellian $f(x) = x^2 G(x, 0, 25)$ where $G(x, \mu, \sigma)$ is the Gauss-function of Eq. 1. For explanation see text.*

The most commonly used measure of the width of a distribution is the variance

$$\sigma^2 \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (6)$$

In practice we always have a limited number of measurements, and thus can only approximate the average and variance for a series of observations, as follows. For the average we take

$$\mu \simeq \bar{x} \equiv \frac{1}{N} \sum_{i=1}^N x_i \quad (7)$$

and for the variance

$$\sigma^2 \simeq s^2 \equiv \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{1}{N-1} \left(\sum_{i=1}^N x_i^2 - N\bar{x}^2 \right) \quad (8)$$

In estimating the average with Eq. 7 we assume that all measurements x_i are independent. In estimating the variance with Eq. 8 we note that we already have used the x_i values to estimate the average, which implies that we have $N - 1$ independent measurements left. (If we know $N - 1$ values of x_i and \bar{x} , we can determine the one remaining x_i .) This is the reason that the sum of $(x_i - \bar{x})^2$ is divided by $N - 1$ (instead of by N). Indeed, if we only have one measurement, i.e. $N = 1$, then the best value for the average is given by this one measurement, but we have no measure for the variance.

Take \bar{x} as a variable in Eq. 8; to find the value of \bar{x} for which s^2 is minimal, we set the derivative of s^2 with respect to \bar{x} to zero:

$$\frac{\partial s^2}{\partial \bar{x}} = \frac{-2}{N-1} \sum_{i=1}^N (x_i - \bar{x}) = \frac{-2}{N-1} \left(\sum_{i=1}^N x_i - N\bar{x} \right) = 0 \quad (9)$$

From the last two terms we obtain once more Eq. 7. This shows that the definition Eq. 7 for the sample average minimizes the sample variance given by Eq. 8.

When the measurements are binned, as is always the case if we have digital measurements but also often in other cases, we find the average and its variance from the binned data. We write the number of bins as B , and denote the value of x in bin b as x_b and the number of measurements in that bin as N_b , where $b = 1, \dots, B$. We normalize the number of measurements in each bin on the total number of measurements, to find the probability that a measurement from the sample falls in bin b as $P_b \equiv N_b / \sum_{b=1}^B N_b$, and rewrite Eqs. 7,8 as

$$\bar{x} = \sum_{b=1}^B P_b x_b \quad (10)$$

and

$$s^2 = \frac{N}{N-1} \left(\sum_{b=1}^B P_b x_b^2 - \bar{x}^2 \right) \quad (11)$$

Higher order moments of distributions with points far from the average ('outliers') are very sensitive to outliers: a large $x_i - \bar{x}$ value dominates much more in the distribution of the $(x_i - \bar{x})^2$'s than in the distribution of $|x_i - \bar{x}|$'s.

For an analogous reason, it is strongly recommended never to use even higher moments, such as the third order moment or *skewness* and the fourth order moment or *kurtosis*. For completeness, here are the definitions of these moments:

$$\text{Skewness} \equiv \frac{1}{N\sigma^3} \sum (x_i - \mu)^3 \simeq \frac{1}{Ns^3} \sum (x_i - \bar{x})^3 \quad (12)$$

and

$$\text{Kurtosis} \equiv \frac{1}{N\sigma^4} \sum (x_i - \mu)^4 - 3 \simeq \frac{1}{Ns^4} \sum (x_i - \bar{x})^4 - 3 \quad (13)$$

where σ is the standard deviation, defined in Eq.6. The subtraction of 3 in the kurtosis is made so that the kurtosis of a Gauss distribution is zero.

In this lecture we will stick to the common practice of mainly using the variance.

1.2 Common distributions

In this section we discuss distributions which often occur or are postulated to occur in nature, or in measurement series. We briefly discuss computer programs to compute these distributions.

1.2.1 The binomial and Poisson distributions

An observer tries to detect photons during a certain time interval. The expected number of photons to arrive in this interval is μ . How many photons will the observer detect? To answer this question, we start by dividing the interval in n equal sub-intervals. In each sub-interval the probability that a photon arrives is thus $p = \mu/n$ (a correction to this statement follows below!); the probability that no photon arrives is $1-p$. The measurement of the observer can thus be considered as a series of n trials to find a photon, each trial having probability p of succeeding. The probability that k photons will be detected is thus given by the binomial function

$$P_B(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k} \quad (14)$$

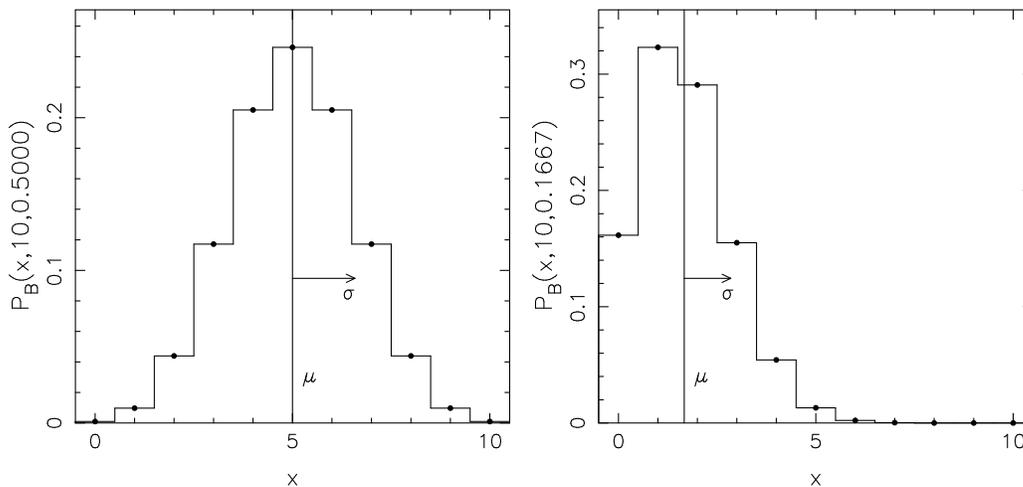


Figure 4: *Binomial distributions for $n = 10$, $p = 0.5$ (left) and $n = 10$, $p = 1/6$ (right).*

The binomial function also describes processes like flipping a coin, or throwing dice. For example, when we flip a coin the probability of getting head is $p = 1/2$, and the probability of getting k times head when we flip the coin 10 times is given by $P_B(k, 10, 0.5)$. When we throw 10 dice, the probability of getting k times a 1 is given by $P_B(k, 10, 1/6)$. These examples are illustrated in Figure 4.

We now return to the experimenter measuring the arrival of photons. If the sub-interval is large, there is a finite probability that more than 1 photon will arrive in it, which is not taken into account in the above argument. To suppress this probability, we take the limit of letting the number of trials n go to infinity, keeping the product $np \equiv \mu$ constant. In this limit the binomial distribution changes into the Poisson distribution, in which the probability of obtaining k photons is given by

$$P_P(k, \mu) = \frac{\mu^k}{k!} e^{-\mu} \quad (15)$$

The derivation is given in the intermezzo. The exponent makes sure that the sum $\sum P_P(k, \mu)$ over all probabilities from $k = 0$ to $k = \infty$ is 1 (remember $e^x = 1 + x + x^2/2! + x^3/3! \dots$).

Our observer does an infinite number of observations (this is a thought experiment...), and should then on average detect μ photons: $\bar{k} = \mu$. The variance $\sigma^2 \equiv \overline{(k - \mu)^2}$ also equals μ . The derivations are given in the intermezzo. Two examples of the Poisson distribution are given in Figure 5.

Intermezzo: derivations

From De Moivre or binomial to Poisson:

$$\begin{aligned} P(k) &= \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k} \\ &= \frac{\mu^k}{k!} \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} \frac{(1 - \mu/n)^n}{(1 - \mu/n)^k} = \frac{\mu^k}{k!} 1 \frac{e^{-\mu}}{1} \end{aligned} \quad (16)$$

Average and variance of Poisson distribution:

$$1 = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} \quad (17)$$

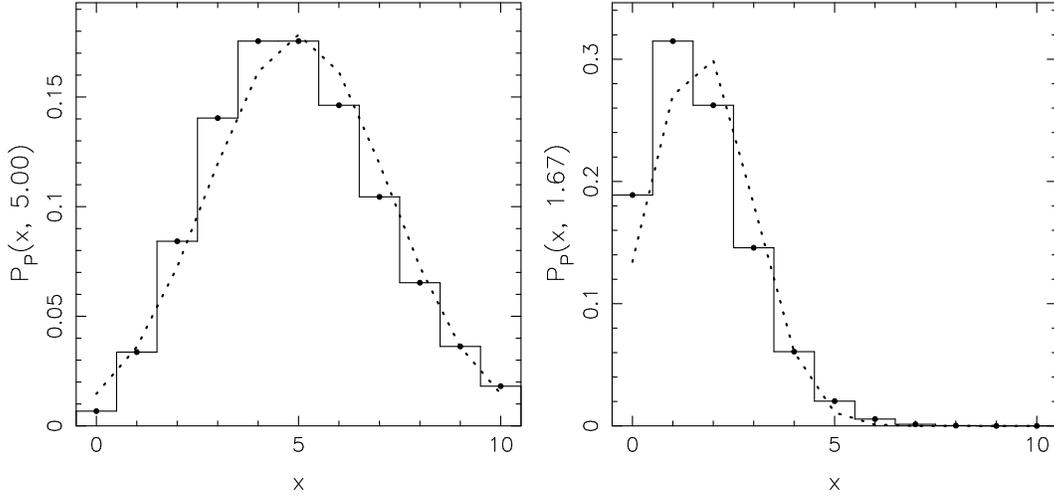


Figure 5: *Poisson distributions for $\mu = 10$ (left) and $p = 10./6$ (right). The dashed lines show Gaussian approximations to the Poisson function: these underestimate the probability of large deviations.*

$$\Rightarrow \frac{\partial}{\partial \mu} 1 = 0 = \sum_{k=0}^{\infty} k \frac{\mu^{k-1}}{k!} e^{-\mu} - \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} = \frac{1}{\mu} \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} e^{-\mu} - 1 = \frac{1}{\mu} \bar{k} - 1 = 0 \Rightarrow \bar{k} = \mu$$

$$\frac{\partial}{\partial \mu} \mu = \frac{\partial}{\partial \mu} \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} e^{-\mu} \Rightarrow 1 = \sum_{k=0}^{\infty} \left(k^2 \frac{\mu^{k-1}}{k!} e^{-\mu} - k \frac{\mu^k}{k!} e^{-\mu} \right) = \frac{1}{\mu} \sum_{k=0}^{\infty} k^2 \frac{\mu^k}{k!} e^{-\mu} - \mu \quad (18)$$

$$\Rightarrow \mu^2 + \mu = \bar{k}^2$$

hence

$$\sigma^2 \equiv \overline{(k - \mu)^2} = \sum_{k=0}^{\infty} (k - \mu)^2 P(k) = \sum_{k=0}^{\infty} k^2 P(k) - 2\mu \sum_{k=0}^{\infty} k P(k) + \mu^2 = \sum_{k=0}^{\infty} k^2 P(k) - \mu^2 = \mu$$

Alternative derivations:

$$\begin{aligned} \bar{k} &= \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} (0 + \mu + 2 \frac{\mu^2}{2!} + \dots) = e^{-\mu} (\mu + \frac{\mu^2}{1!} + \frac{\mu^3}{2!} \dots) \quad (19) \\ &= e^{-\mu} \mu (1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \dots) = e^{-\mu} \mu e^{\mu} = \mu \end{aligned}$$

$$\begin{aligned} \sum_{k=0}^{\infty} k^2 \frac{\mu^k}{k!} &= 0 + \mu + 2\mu^2 + 3 \frac{\mu^3}{2!} \dots = \mu (1 + \frac{\mu}{1!} + 3 \frac{\mu^2}{2!} \dots) = \mu (1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \dots \frac{\mu}{1!} + 2 \frac{\mu^2}{2!} + \dots) \quad (20) \\ &= \mu (e^{\mu} + \mu [1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \dots]) = \mu e^{\mu} + \mu^2 e^{\mu} \end{aligned}$$

hence

$$\sigma^2 \equiv \overline{(k - \mu)^2} = \sum_{k=0}^{\infty} k^2 P(k) - \mu^2 = e^{-\mu} \sum_{k=0}^{\infty} k^2 \frac{\mu^k}{k!} - \mu^2 = \mu + \mu^2 - \mu^2 = \mu$$

1.2.2 The Normal and Gaussian distributions

If μ is large, the Poisson distribution is well approximated by the normal distribution

$$P_G(x) dx = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} x^2 \right] dx \quad (21)$$

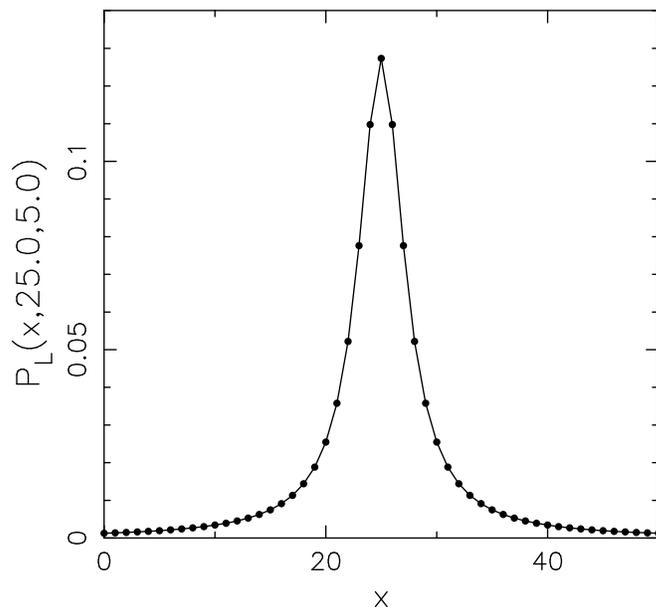


Figure 6: *The Lorentzian distribution for $\mu = 25$ and $\Gamma = 5$.*

We added a finite interval dx here, because in a continuous distribution the probability of getting *exactly* x (i.e. with an infinite number of decimals...) is zero. The more general Gaussian distribution

$$P_G(x, \mu, \sigma)dx \equiv \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}dx \quad (22)$$

is easily converted into the normal distribution Eq. 21 by applying the transformation $z = (x - \mu)/\sigma$. The standard normal distribution is the same as the Gaussian distribution for $\mu = 0$, $\sigma = 1$. The Gauss and normal function have the computational advantage over both the binomial and the Poisson distribution that it is easy to compute; it does not contain such expensive calculations as the factorials. It also describes (or is assumed to describe...) the distribution of measurements in many physical experiments.

1.2.3 The Lorentzian distribution

Spectral lines, or resonance reactions, can often be described with the Lorentzian function

$$P_L(x, \mu, \Gamma) = \frac{1}{\pi} \frac{\Gamma/2}{(x - \mu)^2 + (\Gamma/2)^2} \quad (23)$$

This is a symmetric function around the average μ with a half-width Γ . The integral of the Lorentzian for $-\infty$ to $+\infty$ is 1, as can be verified using the coordinate transformation $z = 2(x - \mu)/\Gamma$:

$$\int_{-\infty}^{\infty} P_L(x, \mu, \Gamma)dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dz}{1 + z^2} = 1 \quad (24)$$

The Lorentzian has very broad wings. As a result, the integral

$$\int_{-\infty}^{\infty} (x - \mu)^2 P_L(x, \mu, \Gamma)dx = \frac{1}{\pi} \frac{\Gamma^2}{4} \int_{-\infty}^{\infty} \frac{z^2 dz}{1 + z^2} \quad (25)$$

is unbounded.

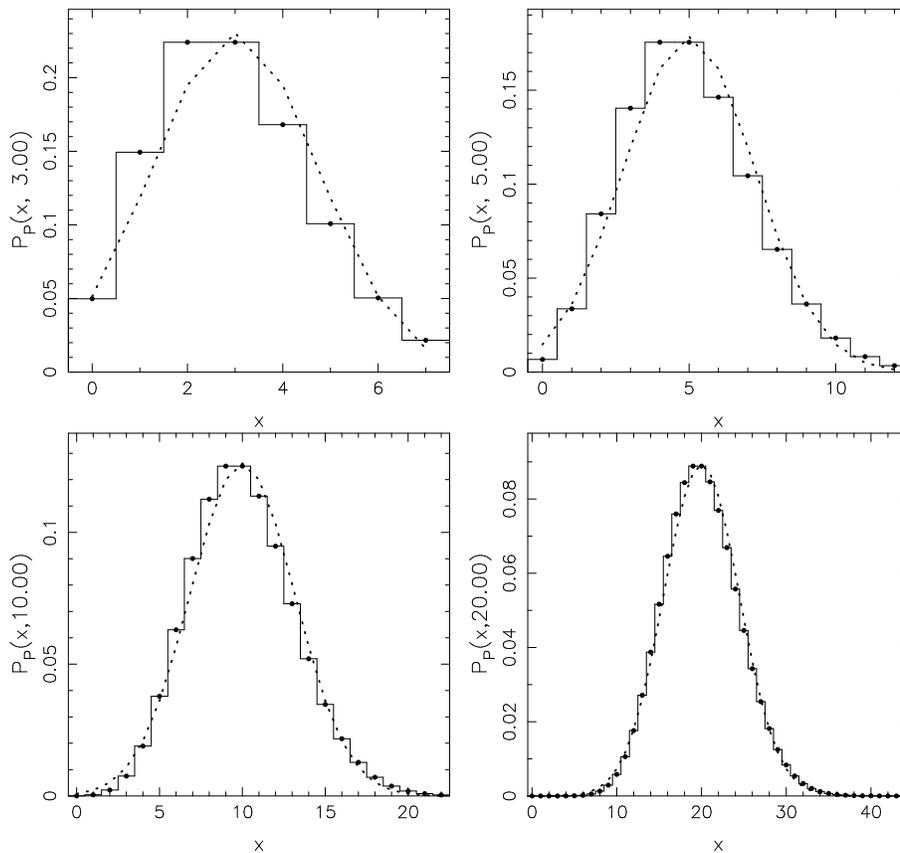


Figure 7: Comparison between the Poisson (solid) and gaussian (dashed) distributions for increasing values of the average. The similarity increases with the average.

1.3 Cumulative probabilities

If we wish to compute the probability of a certain measurement, we must usually compute a cumulative probability. The following argument may illustrate this. Suppose that someone throws a die 3000 times, and obtains a six 500 times. This is obviously an allowed result for a good die. However, if we compute the probability that *exactly* 500 sixes are obtained, we will obtain a low probability (about 2%), as this does not take into account the spread of about $\sqrt{500}$ around the expected number of 500 sixes we may obtain in any single trial. A much better measure for the probability is therefore to compute the probability of getting ' k or less' if we think the obtained number k is too small; or ' k or more' if we suspect the obtained number k is too big.

First we note that the factorials, necessary in calculations with binomial and Poisson function, are closely related to the *gamma function*

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt \quad (26)$$

For integer z , the gamma function equals the factorial, with an offset of one in the argument:

$$n! = \Gamma(n + 1) \quad (27)$$

For small values of n , the factorial can be computed directly as $n! = 1 \times 2 \times 3 \dots n$; for large values we can use the gamma-function. For large n there is the danger that

the factorial is larger than the largest number allowed by the computer language that you use. In that case it is better to calculate the logarithm of the factorial (or equivalently, the logarithm of the gamma function).

The cumulative binomial probability that a binomial process leads to k or more in n trials with probability p is given by the incomplete beta function $I_x(a, b)$:

$$P_B(\geq k, n, p) \equiv \sum_{j=k}^n \binom{n}{j} p^j (1-p)^{n-j} = I_p(k, n-k+1) \quad (28)$$

where

$$I_x(a, b) \equiv \left(\int_0^x t^{a-1} (1-t)^{b-1} dt \right) / \int_0^1 t^{a-1} (1-t)^{b-1} dt \quad (29)$$

The cumulative Poisson probability that a Poisson process with expected value x will lead to a result $k-1$ or less is given by the incomplete Gamma function $Q(a, x)$:

$$P_P(< k, x) \equiv \sum_{j=0}^{k-1} P_P(j, x) = Q(k, x) \equiv \frac{1}{\Gamma(a)} \int_x^\infty t^{a-1} e^{-t} dt \quad (30)$$

The complement of this function is also called the incomplete gamma function (which is somewhat confusing...):

$$P(a, x) \equiv 1 - Q(a, x) \equiv \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt \quad (31)$$

Finally, **the cumulative probability of the Gauss function** (see Eq. 1), can be computed from the *error function*:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (32)$$

Equally useful is the *complementary error function*

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \quad (33)$$

The error functions are given by the incomplete gamma functions:

$$\operatorname{erf}(x) = P(1/2, x^2) \quad (x \geq 0)$$

$$\operatorname{erfc}(x) = Q(1/2, x^2) \quad (x \geq 0)$$

1.4 Remarks on the use of the gaussian

Error distributions are never pure gaussians. However, if a measurement involves a number of processes, each with its own error distribution, the resulting overall error distribution will approximate a gaussian. This theorem is called the central limit theorem. Indeed, measurement error distributions may often be approximated with a gaussian. It never hurts to check whether the errors of your data have a gaussian distribution... It is to be expected that the gaussian approximation is best near the average, and becomes worse the further one moves from the average. In other words, the gaussian gives a correct value at 3σ , say, but cannot be trusted in

detail at $10\text{-}\sigma$. This implies that the difference between, for example, probabilities of 0.01 and 0.001 is significant; but the difference between, for example, 10^{-20} and 10^{-30} usually is not. In general, very small probabilities should not be considered exact.

Many measurements produce integer numbers. This is the case for example if one measures the number of people as a function of height, and also if one measures with a CCD the number of photons as a function of position on the detector. If the numbers are large enough, their error distributions may be approximated with gaussians; however, for small numbers this approximation is bad, and one should use the Poisson distribution. This is illustrated in Fig. 7. For large deviations from the average, one should also use the Poisson distribution: in general, the probability for large deviations is higher in the Poisson distribution than in the gaussian distribution.

As a rule of thumb, one may use the gaussian approximation of the Poisson function if the expectation value is higher than 20. Even in this case, there is a systematic difference between the results obtained with the two distributions. For this reason, and because computers are fast enough to compute Poisson probabilities, it becomes more common to use the Poisson distribution also for large numbers.

If the errors are distributed according to a Gaussian Eq. 22, we can compute how often it happens that the measurement is more than 1, 2 or 3 σ from the expected value. The fraction of values more than $n\sigma$ above the average value is given by

$$\int_{n\sigma+\mu}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} dx = \frac{1}{\sqrt{\pi}} \int_{n/\sqrt{2}}^{\infty} e^{-t^2} dt \quad (34)$$

where we made a transformation

$$\frac{x-\mu}{\sqrt{2}\sigma} \rightarrow t \Rightarrow x = \mu + n\sigma \rightarrow t = \frac{n}{\sqrt{2}} \quad (35)$$

Similarly, we may compute the fraction below $-n\sigma$ from the average. This is the same as the fraction above $n\sigma$ from the average, as follows from the symmetry of the problem, and also from an explicit transformation analogous to the one just made. Thus the total fraction more than $n\sigma$, above or below, from the average value is given by the error function Eq. 33. Some values for these fractions are given in Table 1, together with the n -values for fractions of 90, 99 and 99.9 %.

1.5 Some analytical examples

1.5.1 Average of two measurements

As a first example we consider two measurements N_1 and N_2 . Assuming that both are drawn from a Poisson distribution with the same m , we want to compute the most probable value for m . From Eq. 15 we obtain the combined probability as

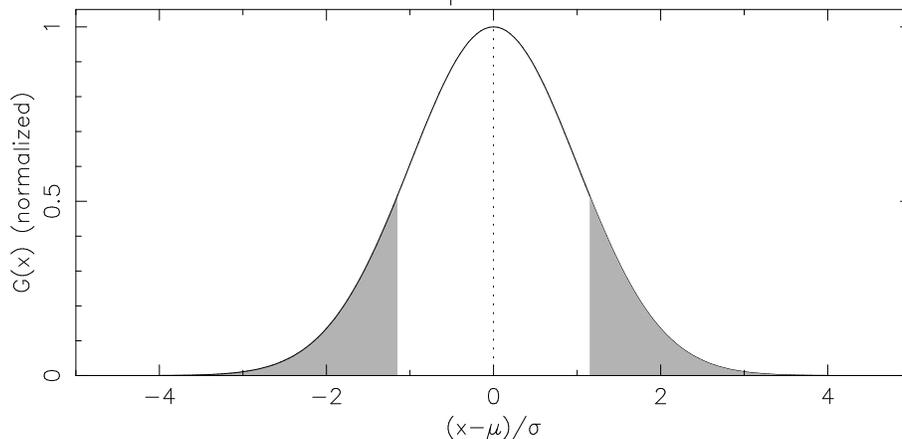
$$L = \frac{m^{N_1} e^{-m}}{N_1!} \frac{m^{N_2} e^{-m}}{N_2!} \quad (36)$$

To find the most probable m we set the derivative of L with respect to m equal to zero, and for computational simplicity we first take the logarithm:

$$\frac{d \ln L}{dm} = 0 = \frac{N_1 + N_2}{m} - 2 \Rightarrow m = \frac{N_1 + N_2}{2} \quad (37)$$

Table 1: *Fractions for the Gauss function between the average and $\pm n\sigma$ (in) and outside this range (out)*

n	in	out	n	in	out
1	0.6827	0.3173	1.645	0.90	0.10
2	0.9545	0.0455	2.575	0.99	0.01
3	0.9973	0.0027	3.890	0.999	0.001



The answer is not unexpected, but nonetheless warrants a closer look, which we will take in Sect. 4.2.1.

1.5.2 Birth ratios of boys and girls

One of the earliest interests of Laplace in statistics is related to the observations that more boys are born than girls. (This is described in Stigler, 1986, pp. 134 sqq., from which we take the observed numbers for the births.) We will not follow Laplace in tackling this, but a modern description.

a. Between 1745 and 1770 in Paris 251 527 boys were born and 241 945 girls. Is the difference significant? To decide this, we assume that the probability for a boy to be born is $p = 0.5$. Then the probability to have $N_1=251\,527$ boys *or more* on a total of $N_P=493\,472$ births is (from Eq. 14)

$$P = \sum_{j=N_1}^{N_P} \binom{N_P}{j} 0.5^{N_P}$$

This sum of binomial probabilities can be computed with the incomplete beta function (for which see *Numerical Recipes*, Sect.6.4, also the very end of that section). Entering the values for N_1 and N_P (N.B.: the routine expects these numbers as real numbers!) we find a probability of 1.1×10^{-42} . We may safely conclude that the probability for a boy is larger than 0.5.

b. In London 737 629 boys and 698 958 girls were born from 1664 to 1757. Is the ratio of boys to girls in London, 1.0553, significantly higher than that in Paris, 1.0396? To answer this question we first derive the best guess for the ratio, under the assumption that it is the same in both cities. Write the number of boys born in London as N_2 and the total number of births in London N_L , we can write the combined probability

$$P = \binom{N_P}{N_1} p^{N_1} (1-p)^{N_P-N_1} \binom{N_L}{N_2} p^{N_2} (1-p)^{N_L-N_2}$$

To find the best guess for p , we take the derivative of the (logarithm of the) probability, and put it equal to zero:

$$\frac{d \ln P}{dp} = \frac{N_1 + N_2}{p} - \frac{(N_P + N_L) - (N_1 + N_2)}{1 - p} = 0 \Rightarrow p = \frac{N_1 + N_2}{N_P + N_L}$$

Entering the numbers we find $p = 0.5125$ and with this the probability that the number of boys in Paris is N_1 or higher,

$$P_P = P = \sum_{j=N_1}^{N_P} \binom{N_P}{j} 0.5125^j (1 - 0.5125)^{N_P - j}$$

is almost 1, but the probability for London for N_2 or higher is 0.012. (We compute this again with the incomplete beta function.) Thus the probability that the birth ratio in London is the same as in Paris is about 1%.

1.6 Computing the various distributions

In the bad old days, one was forced to use tables in thick, heavy books whenever one used the distributions described above. Nowadays, fast computers allow us to compute the quantities we need, with relative ease. Chapter 6 of *Numerical recipes* (2nd edition) contains good explanations of these functions. A good introduction for using these functions in *Python* may be found in:

www.codeproject.com/Articles/38402/Getting-started-with-the-SciPy-Scientific-Python-1

1.7 Exercises

Exercise 1. Show analytically that the average and the variance of the Gauss function Eq. 1 are given by μ and σ^2 .

1.8 Computer exercises

Computer Exercise 1. We start with an exercise related to the machine accuracy of the computer. It is adapted from Meeus, *Astronomical Algorithms*² (2000).

Write a code, which first sets $x = 1/3$ (x is a real) and then n times replaces x with $(9x + 1)x - 1$, writing each replacement to the screen. Run the code. Do you understand how the error spreads?

Computer Exercise 2. For this exercise, use the functions available in `numpy`.

a. Write a computer programme which computes the binomial probability for given k , n , p . To check your program, verify the following statement: if one throws 10 dice, the probability that precisely 7 three's are up is 2.48×10^{-4} .

b. Write a computer programme which computes the Poisson probability for given k , μ . To check your program, verify the following statement: if the expected number of arriving counts is 15, the probability of measuring 30 counts is 2.21×10^{-4} , the probability of measuring 0 counts is 3.06×10^{-7} .

c. Write a computer programme which computes the Gauss function for given x , μ , σ . To check your program, verify the following statement: if one estimates the probabilities calculated in b) by approximating the Poisson function with a Gaussian, one finds 5.70×10^{-5} for both.

Computer Exercise 3. Write a code to verify the computations in Sect.1.5.2

Computer Exercise 4. A pulsar, short for *pulsating* radio source, is observed with an X-ray satellite, and 13 photons are detected from the source position. The data are folded on the known pulse period, and binned in 5 bins. It is then found that 8 of the 13 photons fall in one bin.

a. Describe how one can determine what the probability is of finding (exactly) 8 of 13 photons in bin one (of five). Which parent distribution must you use? Write a computer code to compute this probability. Write two versions: one which takes the input numbers from the command line; and one which prompts for the input numbers while executing the code.

b. Argue that the probability that one finds 8 counts in *any* of the five bins is five times the result of a). Nota bene: suppose in a) we had computed the probability of finding 5 or more counts in bin one. In that case the probability of finding 5 or more counts in *any* bin cannot be found by multiplying the probability for one bin with five. Why not?

c. Argue that the correct probability to compute is the cumulative probability of finding 8 *or more* counts in any of the five bins. Extend your computer code to compute this probability, by explicitly summing from 8 to 13 (inclusive). The correct answer is 0.006. Then, use the function for the incomplete beta-function in `scipy.special`

This Exercise is based on an observation with the ROSAT (Roentgen Satellite) of a pulsar in the globular star cluster M 28. The original authors claimed a probability of about 3×10^{-4} that the 8 photons are in one bin due to chance. It turns out that the 13 photons are not from the pulsar, so that the apparent pulsation of the X-ray photons is really due to chance! (Verbunt, 2001, *Astronomy and Astrophysics* 368, 137; sect. 3.44.)

1.9 What is wrong?

The Moon causes tides in the oceans on Earth, and because of friction in these tides, angular momentum is transferred from the rotation of the Earth to the orbit of the Moon. This implies that the rotation of the Earth slows down (the length of the day increases by about 0.002 s/century), and the distance to the Moon increases by about 3.84 cm/yr).

Can one measure the length of the day in the distant past? In 1973 it was suggested that one can use the variation of the thickness in layers of chalk in some fossil shells. The idea was that a shell adds a layer every day, and that the thickness of the layer depends on the amount of light, which changes with the season of the year. Thus, by carefully counting the number of layers one can determine the number of days per year. Since the length of the year is (virtually) constant, one thereby determines the length of the day. I have taken the values with error estimates from a review on this topic, and show them in Figure 8, together with a theoretical curve (dotted line). From this figure it is immediately clear that the ‘observed’ numbers have been doctored. Explain how you can see this.

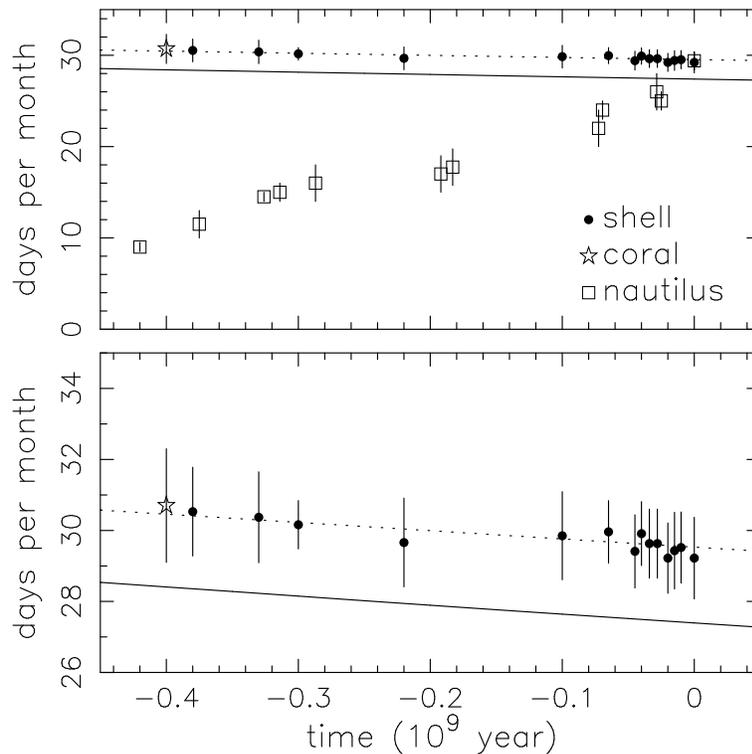


Figure 8: *The number of days per month and per year as determined from counts of chalk layers in fossil shells. The dashed lines gives the theoretical curve, the dots the measurements with 1- σ error bars.*

For more detail see: <http://www.astro.ru.nl/~fverbunt/binaries/earth.pdf>

2 Error propagation

Suppose we have a function f of a number of variables u, v, \dots :

$$f \equiv f(u, v, \dots) \quad (38)$$

We want to know how we can estimate the variance of f ,

$$\sigma_f^2 \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (f_i - \bar{f})^2 \quad (39)$$

from the variances $\sigma_u, \sigma_v, \dots$ of the variables u, v, \dots . To derive this we will make the assumption, which is usually only approximately correct, that the average of f is well approximated by the value of f for the averages of the variables:

$$\bar{f} = f(\bar{u}, \bar{v}, \dots) \quad (40)$$

We make a Rayleigh-Taylor expansion of f around the average:

$$f_i - \bar{f} \simeq (u_i - \bar{u}) \frac{\partial f}{\partial u} + (v_i - \bar{v}) \frac{\partial f}{\partial v} + \dots \quad (41)$$

to write the variance in f as

$$\begin{aligned} \sigma_f^2 &\simeq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left[(u_i - \bar{u}) \frac{\partial f}{\partial u} + (v_i - \bar{v}) \frac{\partial f}{\partial v} + \dots \right]^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left[(u_i - \bar{u})^2 \left(\frac{\partial f}{\partial u} \right)^2 + (v_i - \bar{v})^2 \left(\frac{\partial f}{\partial v} \right)^2 + 2(u_i - \bar{u})(v_i - \bar{v}) \frac{\partial f}{\partial u} \frac{\partial f}{\partial v} + \dots \right] \end{aligned} \quad (42)$$

We rewrite this by using the definitions for the variances of u and v ,

$$\sigma_u^2 \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (u_i - \bar{u})^2; \quad \sigma_v^2 \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (v_i - \bar{v})^2 \quad (43)$$

and defining the covariance of u and v as

$$\sigma_{uv}^2 \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (u_i - \bar{u})(v_i - \bar{v}) \quad (44)$$

to obtain:

$$\sigma_f^2 = \sigma_u^2 \left(\frac{\partial f}{\partial u} \right)^2 + \sigma_v^2 \left(\frac{\partial f}{\partial v} \right)^2 + 2\sigma_{uv}^2 \frac{\partial f}{\partial u} \frac{\partial f}{\partial v} + \dots \quad (45)$$

If the differences $u_i - \bar{u}$ and $v_i - \bar{v}$ are not correlated, the sign of their product is as often positive as negative, and the sum in eq. 44 will be small as subsequent terms (almost) cancel. On the other hand, if the differences are correlated, most products $(u_i - \bar{u})(v_i - \bar{v})$ will be positive, and the cross-correlation term in Eq. 45 can be large as subsequent terms add up.

2.1 Examples of error propagation

2.1.1 weighted sum: $f = au + bv$

For the partial derivatives we have

$$\frac{\partial f}{\partial u} = a, \quad \frac{\partial f}{\partial v} = b \quad (46)$$

and substituting these in eq. 45 we obtain

$$\sigma_f^2 = a^2\sigma_u^2 + b^2\sigma_v^2 + 2ab\sigma_{uv}^2 \quad (47)$$

Note that a and b can each have positive or negative sign. Their signs only have an effect on the cross-correlation term, which as a consequence can be negative, and *make the variance smaller!* For example, if each u_i is accompanied by a v_i such that $v_i - \bar{v} = -(b/a)(u_i - \bar{u})$, then $f = a\bar{u} + b\bar{v}$ for all u_i, v_i pairs, and $\sigma_f^2 = 0$.

2.1.2 product: $f = auv$

With

$$\frac{\partial f}{\partial u} = av, \quad \frac{\partial f}{\partial v} = au \quad (48)$$

we obtain

$$\frac{\sigma_f^2}{f^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + \frac{2\sigma_{uv}^2}{uv} \quad (49)$$

2.1.3 division: $f = au/v$

$$\frac{\partial f}{\partial u} = \frac{a}{v}, \quad \frac{\partial f}{\partial v} = -\frac{au}{v^2} \quad (50)$$

hence

$$\frac{\sigma_f^2}{f^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} - \frac{2\sigma_{uv}^2}{uv} \quad (51)$$

2.1.4 exponent: $f = ae^{bu}$

$$\frac{\partial f}{\partial u} = bf \quad (52)$$

hence

$$\frac{\sigma_f}{f} = b\sigma_u \quad (53)$$

2.1.5 power: $f = au^b$

$$\frac{\partial f}{\partial u} = \frac{bf}{u} \quad (54)$$

hence

$$\frac{\sigma_f}{f} = b\frac{\sigma_u}{u} \quad (55)$$

Exercise 2. $x = a \ln(\pm bu)$. Give the relation between σ_x and σ_u .

2.2 Exercises

Exercise 3. Radiopulsar PSR J 1713 + 0747 has an observed period $P = 4.5701 \times 10^{-3}$ s and period derivative $\dot{P} = 8.52 \times 10^{-21}$, both with (for the purpose of this problem) negligible errors. Two components of its proper motion μ have been measured as $\mu_\alpha = 4.9 \pm 0.3$ mas/yr and $\mu_\delta = -4.1 \pm 1.0$ mas/yr; its distance is estimated from the dispersion measure as $d = 1.1 \pm 0.3$ kpc.

As noted by Shklovskii, the ratio \dot{P}/P must be corrected for the apparent acceleration due to the proper motion, so that the intrinsic ratio is:

$$\left(\frac{\dot{P}}{P}\right)_i = \frac{\dot{P}}{P} + \frac{\mu^2 d}{c} \quad (56)$$

The (intrinsic) spindown age of a radio pulsar is

$$\tau_{c,i} = \left(\frac{P}{2\dot{P}}\right)_i \quad (57)$$

Calculate the intrinsic spindown age and its error.

3 Errors distributed as a Gaussian and the least squares method

Suppose we have a series of measurements y_i with associated errors distributed according to a Gaussian with width σ_i , i.e. each measurement is drawn from a Gaussian with width σ_i around a model value y_m . For one measurement, the probability $P(y_i) \equiv P_i$ of obtaining a measurement y_i , in an interval Δy , is then given by

$$P_i \Delta y = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(y_i - y_m)^2}{2\sigma_i^2}} \Delta y \quad (58)$$

where we allow that different measurements have different associated errors σ_i . If we have a series of N measurements the overall probability of obtaining such a series is:

$$P(\Delta y)^N \equiv \prod_{i=1}^N (P_i \Delta y) = \frac{1}{(2\pi)^{N/2} \prod_i \sigma_i} \exp \left[-\frac{1}{2} \sum_{i=1}^N \frac{(y_i - y_m)^2}{\sigma_i^2} \right] \Delta y^N \quad (59)$$

The highest probability P is that for which

$$\chi^2 \equiv \sum_{i=1}^N \chi_i^2 \equiv \sum_{i=1}^N \frac{(y_i - y_m)^2}{\sigma_i^2} \quad (60)$$

is smallest. If we wish to determine the most probable model value for y_m for a series of measurements y_i , we thus must find the value(s) for y_m for which the sum of the squares $(y_i - y_m)^2/\sigma_i^2$ is minimal. This is called the method of least squares and was described first in 1804 by Gauss, following earlier work by Lagrange and Laplace.

According to the assumption that the errors are gaussian, as stipulated above, each χ_i is a random draw from a normal distribution; the sum of the χ_i^2 is called chi-square. If we have N measurements, and we fit M parameters, we have $N - M$ independent measurements left. The probability distribution for χ^2 is called the chi-square distribution for $N - M$ degrees of freedom. This distribution is what we get if would draw $N - M$ random samples from a normal distribution, square them, and add the squares, and repeat this many times, each time producing one χ^2 . The probability of a given χ^2 can be computed with the incomplete gamma-function, defined in Eq.30, as follows: the probability that an observed χ_{obs}^2 or greater is reached for $N - M$ degrees of freedom is:

$$P(\chi^2 \geq \chi_{obs}^2) = Q \left(\frac{N - M}{2}, \frac{\chi_{obs}^2}{2} \right) \quad (61)$$

If the probability that we get is very small, then something is wrong. The most obvious possibility is that the model is wrong. Another possibility is that the errors have been under-estimated, so that we demand a better agreement with the model than the measurement accuracy warrants. A third possibility is that the errors are not distributed as Gaussians. This latter possibility tends to increase the observed χ^2 ; and often causes us to accept probabilities somewhat lower than one would superficially expect. Thus in many cases, a probability of 0.05 is considered quite acceptable. The reason one can do this, is that really wrong models easily produce

much smaller probabilities, less than 0.000001, say. Also, even in a perfect statistical process, a probability of 5% arises on average once every 20 trials, so that finding such a probability due to chance is quite common. If one finds in a certain experiment that one always produces low probabilities, one must look into it and try to determine the cause.

It is worth noting that apparently significant results can arise from ignoring negative results. The most obvious case is the winning of a lottery. Suppose the winner is the person who has guessed the correct integer number less or equal to one million. The probability for one person to have the winning number is one in a million, but if several million people participate we expect several to have guessed the correct number. A less obvious case is an experiment which is repeated many times. A well-known case here is the stock-broker. Suppose that one starts with ten million people, who predict how stocks change in value. After a year, select the people who are among the ten per cent best predictors; repeat this five times (so that six rounds have been made), then there will be (on average) ten stock brokers who have predicted among the best 10% for six years in a row, if the prediction process is purely random! It is thus by no means obvious that a stock-broker who has predicted correctly for many years in a row is better than random – one can only find out if one knows how many stock-brokers there are. . . . (Indeed, in experiments in which a gorilla selects the stocks, the gorilla often does better than the professional stock broker firms.)

When we make a fit of a model to a dataset, the result should always consist of three parts:

1. the best values of the parameters a_1, a_2, \dots
2. the errors on these parameter values
3. the probability that the measured χ^2 is obtained by chance; i.e. the probability that the model adequately describes the measurements (Eq. 61)

It is very important always to provide all these three desiderata.

We will discuss three cases for y_m , viz. i) a constant, in which y_m is the same for all i ; ii) a straight line $y_m = a + bx$ where y_m depends on the variable x and the model parameters a and b ; and iii) the general case in which y_m depends on a variable x and a set of parameters a_1, a_2, \dots , which we write as $y_m = y_m(x; \vec{a})$.

3.1 Weighted averages

If the model for is $y_m = a$, with a a constant. The least squares method tells us that we find the most probable value for a by minimizing eq. 60 with respect to a , i.e.

$$\frac{\partial}{\partial a} \left[\sum_{i=1}^N \frac{(y_i - a)^2}{\sigma_i^2} \right] = 0 \Rightarrow \sum_{i=1}^N \frac{y_i - a}{\sigma_i^2} = 0 \quad (62)$$

Thus the most probable value for a is given by

$$a = \frac{\sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2}} \quad (63)$$

and a is called the weighted average of y_i .

To obtain an estimate of the error in a , we note that it is a function of the variables y_1, y_2, \dots . If the measurements y_i are not correlated, we find the variance for a from (see eq. 45):

$$\sigma_a^2 = \sum_{i=1}^N \left[\sigma_i^2 \left(\frac{\partial \bar{y}}{\partial y_i} \right)^2 \right] = \sum_{i=1}^N \left[\sigma_i^2 \left(\frac{1/\sigma_i^2}{\sum_{k=1}^N (1/\sigma_k^2)} \right)^2 \right] \quad (64)$$

which gives

$$\sigma_a^2 = \frac{1}{\sum_{i=1}^N (1/\sigma_i^2)} \quad (65)$$

In the case where all measurement errors are identical, $\sigma_i \equiv \sigma$, eqs. 63,65 can be further simplified to

$$a = \frac{(1/\sigma^2) \sum_{i=1}^N y_i}{(1/\sigma^2) \sum_{i=1}^N (1)} = \frac{1}{N} \sum_{i=1}^N y_i \quad (66)$$

and

$$\sigma_a^2 = \frac{\sigma^2}{N} \quad (67)$$

3.2 Fitting a straight line

We write the model of a straight line as

$$y_m(x, a, b) = a + bx \Rightarrow y_m(x_i, a, b) = a + bx_i \quad (68)$$

where the parameters to be fitted are a and b . We now wish to minimize the χ^2 , according to Eq. 60, with respect to a and b , and do this by looking for the values for which the derivative of χ^2 is zero:

$$\frac{\partial \sum_{i=1}^N [(y_i - a - bx_i)/\sigma_i]^2}{\partial a} = 0 \Rightarrow \sum_{i=1}^N \left(\frac{y_i - a - bx_i}{\sigma_i^2} \right) = 0 \Rightarrow \sum_{i=1}^N \frac{y_i}{\sigma_i^2} - a \sum_{i=1}^N \frac{1}{\sigma_i^2} - b \sum_{i=1}^N \frac{x_i}{\sigma_i^2} = 0 \quad (69)$$

$$\frac{\partial \sum_{i=1}^N [(y_i - a - bx_i)/\sigma_i]^2}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \frac{x_i(y_i - a - bx_i)}{\sigma_i^2} = 0 \Rightarrow \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2} - a \sum_{i=1}^N \frac{x_i}{\sigma_i^2} - b \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} = 0 \quad (70)$$

The important thing to note now is that all the sums can be evaluated without knowledge of a or b . To make the equations easier to interpret we define the sums as

$$\begin{aligned} \sum_{i=1}^N \frac{1}{\sigma_i^2} &\equiv S; & \sum_{i=1}^N \frac{x_i}{\sigma_i^2} &\equiv S_x; & \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} &\equiv S_{xx} \\ \sum_{i=1}^N \frac{y_i}{\sigma_i^2} &\equiv S_y; & \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2} &\equiv S_{xy}; & \Delta &\equiv SS_{xx} - (S_x)^2 \end{aligned}$$

and rewrite the above two equations as two equations for two unknowns a and b :

$$aS + bS_x - S_y = 0$$

$$aS_x + bS_{xx} - S_{xy} = 0$$

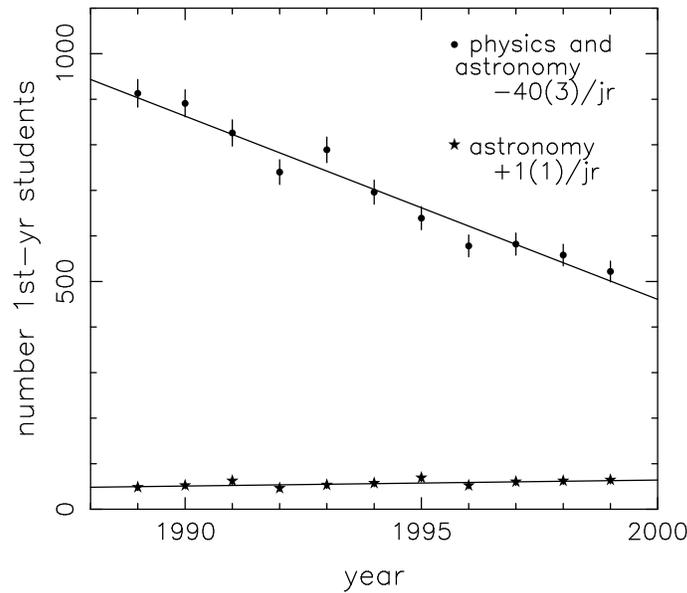


Figure 9: *Total number of 1st-year students in the Netherlands in physics and astronomy (above) and in astronomy only (below) between 1990 and 2000. Assumed poissonian errors are indicated. Data taken from the Nederlands Tijdschrift voor Natuurkunde, which stopped publishing these numbers in 2000. . .*

with the solutions

$$a = \frac{S_{xx}S_y - S_xS_{xy}}{\Delta} \quad (71)$$

$$b = \frac{SS_{xy} - S_xS_y}{\Delta} \quad (72)$$

To find the errors in a and b , we note that a and b depend on the independent parameters y_i , such that

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - x_i S_x}{\sigma_i^2 \Delta}; \quad \frac{\partial b}{\partial y_i} = \frac{x_i S - S_x}{\sigma_i^2 \Delta} \quad (73)$$

and use Eq. 45 to obtain (after some rewriting)

$$\sigma_a^2 = \frac{S_{xx}}{\Delta}; \quad \sigma_b^2 = \frac{S}{\Delta} \quad (74)$$

As the third ingredient of our fit we determine the probability that a good fit would produce the observed χ^2_{obs} or bigger as

$$P(\chi^2 \geq \chi^2_{obs}) = Q\left(\frac{N-2}{2}, \frac{\chi^2_{obs}}{2}\right) \quad (75)$$

As an illustration of some more subtleties we have a look at Figure 9, where the number of new physics and astronomy students in the Netherlands are plotted for the last decade of the 20st century. The first point concerns the errors in measured integer numbers. It is assumed in the figure that the error on the number of students is given by the square root of the number, and these errors are shown. The reader may be surprised now, and say: 'Why is there an error? After all, we know exactly how many students there are in a given year!'. The answer is that *for the purpose of*

fitting a model the numbers of students can vary due to chance, and that the actual number in a year is drawn from a distribution (poissonian in this case) around the expected value. Similarly, if an X-ray experiment detects N photons in a given time interval, in a given energy range, then the error that we assign to this number when we fit it in a model is *not* zero – even though we know exactly how many photons were measured!

The second point concerns a good choice of a and b . If we fit the number of students as $N(t) = a + bt$ where t is the year, then a gives the number of students for the year 0. Apart from the fact that this number is meaningless in the current context, there are two other problems associated with this choice. The first one is that the sums involving x_i -values in Eqs. 71–74 are very large numbers, so that subtracting them from one another (as in computing Δ) easily leads to roundoff errors, and thus to wrong results. The second one is that in this choice of a and b there will be a very strong correlation between the errors: if we change b a little, a changes dramatically in one direction. Both these problems can be circumvented by centering the time interval around the point of fitting, i.e. by effectively fitting $N = a + b(t - 1994)$. Now, the sums involve smaller numbers, and the pivot point around which a and b vary is such that the correlation of their variations are minimized.

Also for non-linear fits, it is good practice in astronomy to define the time with respect to some fiducial point somewhere near the middle of the measurements.

3.3 Non-linear models: Levenberg-Marquardt

When our model involves a parameter a which enters the y_m , and through this the χ_i , non-linearly, we cannot execute a summation without having (an estimate for) the value of a . As an example, consider the function $y = \sin(ax)$, which has (see Eq. 60)

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^N \frac{[y_i - \sin(ax_i)]x_i \cos(ax_i)}{\sigma_i^2} \quad (76)$$

and we see that the summation cannot be executed without a value for a .

In general, this means that a model $y(x, \vec{a})$ which is non-linear in \vec{a} can be fitted to a set of data only iteratively, in the sense that one needs a first set of values for \vec{a} , and then finds successive improvements of these values.

To understand how this is done, we discuss first a one-dimensional case where there is one parameter a that is fitted, so that $\chi^2 = \chi^2(a)$. We are then looking for the value of a for which $\chi^2(a)$ is minimized.

When one is far from the minimum, we can use the derivative $\partial \chi^2 / \partial a$ to decide in which direction to look for the improved value:

$$a_{n+1} = a_n - K \frac{\partial \chi^2}{\partial a} \quad (77)$$

where K is a constant (the value of which we discuss below).

Once we get close to the minimum of χ^2 , this way of stepping towards a better solution becomes less efficient, since near the minimum the first derivative approaches zero. Close to the minimum we approximate χ^2 as a quadratic function of a :

$$\chi^2(a) = \chi^2(a_{min}) + q(a - a_{min})^2; \quad \frac{\partial \chi^2}{\partial a} = 2q(a - a_{min}); \quad \frac{\partial^2 \chi^2}{\partial a^2} = 2q \quad (78)$$

where a_{min} is the parameter value at which χ^2 has its minimum. We combine the derivatives to find

$$(a - a_{min}) \frac{\partial^2 \chi^2}{\partial a^2} = \frac{\partial \chi^2}{\partial a} \Rightarrow a_{n+1} = a_n - \frac{\partial \chi^2 / \partial a}{\partial^2 \chi^2 / \partial a^2} \quad (79)$$

Now make the step to a more-dimensional case, i.e. with more than one parameter, and write for model $y_m = y_m(x, \vec{a})$:

$$\frac{\partial \chi^2}{\partial a_k} = -2 \sum_{i=1}^N \frac{[y_i - y_m]}{\sigma_i^2} \frac{\partial y_m}{\partial a_k} \equiv -2\beta_k \quad (80)$$

and

$$\frac{\partial \chi^2}{\partial a_k \partial a_l} = 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y_m}{\partial a_k} \frac{\partial y_m}{\partial a_l} - [y_i - y_m] \frac{\partial^2 y_m}{\partial a_k \partial a_l} \right] \equiv 2\alpha_{kl} \quad (81)$$

The second term on the right hand side is expected to be small with respect to the first one, because $y_i - y_m$ will almost equally often be positive as negative, and subsequent terms in the summation will almost cancel. To save computing time, one can drop the second term. It often turns out that the iteration to the best solution becomes more stable by doing this.

For the more dimensional case we can write for Eq. 79:

$$\beta_k = \sum_{l=1}^M \alpha_{kl} \Delta a_l; \quad \Delta a_l = a_l - a_{l,min} \quad (82)$$

If the diagonal terms in α are much bigger than the off-diagonal terms, this reduces to the linear case

$$\beta_k = \lambda \alpha_{kk} \Delta a_k \quad (83)$$

which is similar to Eq. 77, when we scale the proportionality constant K with the second derivative, with scaling parameter λ .

The Levenberg-Marquardt method now continues by adding the two last equations:

$$\beta_k = \sum_{l=1}^M \alpha'_{kl} \Delta a_l \quad \text{where} \quad \alpha'_{kl} = \alpha_{kl} \quad (\text{if } k \neq l); \quad \alpha'_{kl} = \alpha_{kl}(1 + \lambda) \quad (\text{if } k = l) \quad (84)$$

For large λ Eq. 84 approaches Eq. 83, and for small λ it approaches Eq. 82.

The solution of the equation then proceeds as follows. One picks an initial solution, and a small value for λ (i.e. one hopes that the solution is already close enough for the quadratic approximation). Compute the χ^2 for the initial solution, and compute a new value for \vec{a} with Eq. 84. If the χ^2 for the new solution is smaller (larger) than for the old one, then the quadratic approach does (doesn't) work, and one should decrease (increase) λ to get closer to the purely quadratic (linear) method of Eq. 82 (83). This process is iterated until the minimum χ^2 is found.

We also need the errors on the best parameters \vec{a} . For this we use the general relation (near minimum)

$$\Delta \chi^2 = \Delta \vec{a} \cdot \vec{\alpha} \cdot \Delta \vec{a} \quad (85)$$

If the best values of the parameters are not correlated, then the matrix α is close to diagonal, i.e. the off-diagonal elements α_{kl} are much smaller than the elements

α_{kk} on the diagonal. In that case the difference distance Δa_k to the best value a_k is given by:

$$\Delta a_k^2 = \frac{\Delta\chi^2}{\alpha_{kk}} \quad (86)$$

and to estimate a 1- σ error in a_k we may enter $\Delta\chi^2 = 1$ in Eq. 86.

When the errors are correlated, the situation is more complex, and the matrix A (more precisely, its inverse matrix $C = A^{-1}$) gives the correlation between the deviations of the parameters from their best values. For details we refer to Numerical Recipes.

3.4 Are the extra parameters significant?

When we have a model with M_A parameters, but suspect that a model with more parameters is necessary, we can make a fit with $M_B > M_A$ parameters. For example, if we describe a spectrum with a continuum, and we think that there may be a spectral line in the spectrum, we can add the line. In the case of a gaussian line profile, this leads to 3 extra parameters: the line strength, the line center, and the line width. With more parameters, the fit obviously is better. But is it significantly better? To decide that, we note that the difference $\chi^2_B - \chi^2_A$ itself follows a χ^2 distribution with $M_B - M_A$ degrees of freedom.

Table 2: Differences in χ^2 required to decide whether the added parameters lead to a significant improvement. The first column gives the required significance level, the following columns the corresponding minimum $\Delta\chi^2$, for 1,2...6 added parameters

P	1	2	3	4	5	6
0.683=1 σ	1.00	2.30	3.53	4.72	5.89	7.04
0.954=2 σ	4.00	6.16	8.02	9.70	11.3	12.8
0.9973=3 σ	9.00	11.8	14.2	16.3	18.2	20.1
0.90	2.71	4.61	6.25	7.78	9.24	10.6
0.99	6.63	9.21	11.3	13.3	15.1	16.8
0.9999	15.1	18.4	21.1	23.5	25.7	27.8

3.5 Computer exercises

In computer exercise 5, we determine the age of the Earth from the abundances of isotopes of rubidium and strontium in meteorites. We first discuss how this method works. The discussion is adapted from Verbunt, *Het Leven van Sterren* (2005, 2010²). In computer exercise 6, we determine the mass of the first detected exoplanet around a solar-type star. For this we first derive the observed radial velocity of a binary star.

3.5.1 The age of the earth from meteorites

The best estimate of the age of the solar system is based on research of meteorites. The age of meteorites can be determined thanks to the decay of long-living radioactive elements. As an example we discuss the decay of a neutron in the core of a

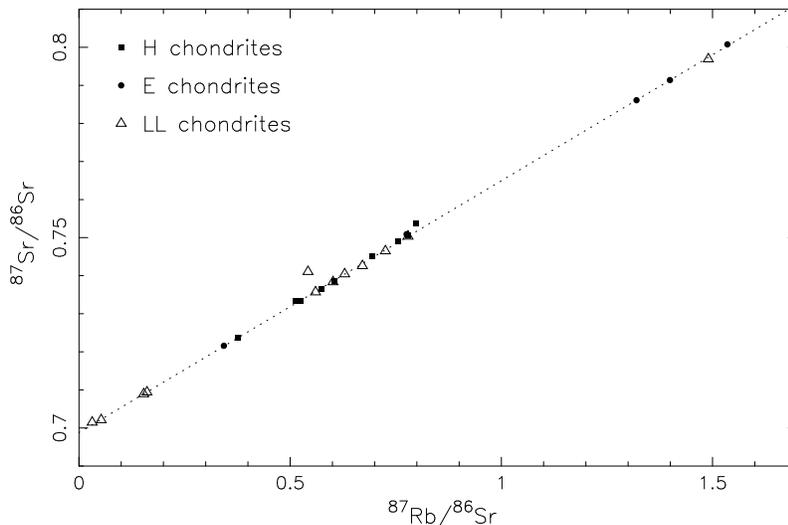


Figure 10: *The ratio $^{87}\text{Sr}(t)/^{86}\text{Sr}$ as a function of $^{87}\text{Rb}(t)/^{86}\text{Sr}$ for various types of meteorites. The dotted line is plotted only to guide the eye.*

rubidium atom into a proton, under emission of an electron and a neutrino, turning rubidium into strontium:



The amount of rubidium at time t is:

$$^{87}\text{Rb}(t) = ^{87}\text{Rb}(0)e^{-\lambda t} \quad (88)$$

where $^{87}\text{Rb}(0)$ is the initial amount of rubidium, and λ the decay constant. The amount of strontium-87 increases accordingly:

$$^{87}\text{Sr}(t) = ^{87}\text{Sr}(0) + (^{87}\text{Rb}(0) - ^{87}\text{Rb}(t)) = ^{87}\text{Sr}(0) + ^{87}\text{Rb}(0)(e^{\lambda t} - 1) \quad (89)$$

We divide this equation by the amount of strontium-86; this amount doesn't change. We obtain:

$$\frac{^{87}\text{Sr}(t)}{^{86}\text{Sr}} = \frac{^{87}\text{Sr}(0)}{^{86}\text{Sr}} + (e^{\lambda t} - 1) \frac{^{87}\text{Rb}(0)}{^{86}\text{Sr}} \quad (90)$$

The meteorites that arrive on Earth are fragments released in collisions between planetoids. The planetoids were (partially) melted soon after their formation, and solidified afterwards. This process partially separates the different elements: strontium solidifies at a higher temperature (769 C) than rubidium (39 C). The initial ratio of rubidium to strontium $^{87}\text{Rb}(0)/^{86}\text{Sr}$ therefore was lower in the relatively rapidly cooling surface layers of the planetoid, and higher in the slowly cooling core. The solidification however does not differentiate between strontium-86 and strontium-87; the initial ratio $^{87}\text{Sr}(0)/^{86}\text{Sr}$ was the same everywhere in the planetoid. The meteorites released from a planetoid therefore cover a range in values for the ratio $^{87}\text{Rb}(t)/^{86}\text{Sr}$.

For a number of meteorites the ratios strontium-87 to strontium-86 and rubidium-87 to strontium-86 have been measured (Figure 10). By fitting the straight line Eq. 90 to the data, we may determine the initial ratio $^{87}\text{Sr}(0)/^{86}\text{Sr}$, and the factor $(e^{\lambda t} - 1)$. As the decay constant is known, we can determine the age of the meteorites. With $\lambda = 1.42 \times 10^{-11} \text{ yr}^{-1}$ we find an age $t = (4.51 \pm 0.01) \times 10^9 \text{ yr}$.

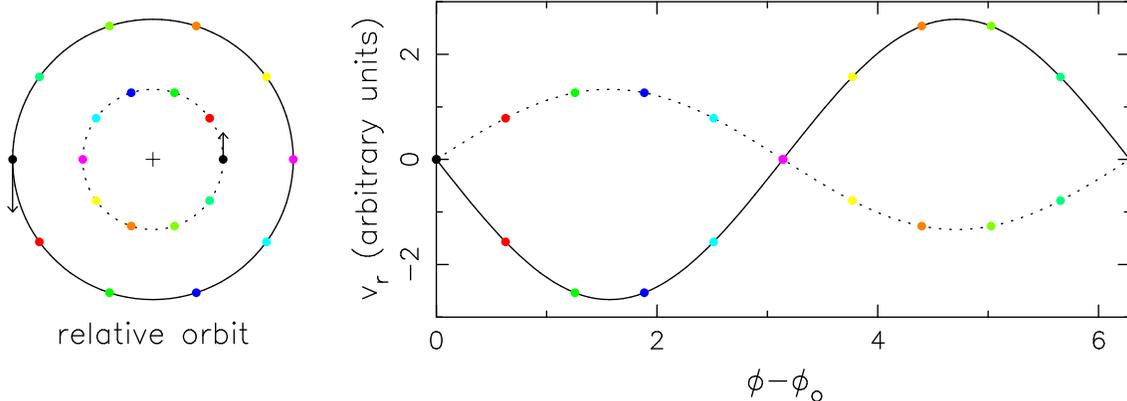


Figure 11: (left): The orbits of two stars, with mass ratio 2, around the center of gravity (indicated with +). The colored dots indicates some positions in the orbits, equal colours for simultaneous positions. (right): the radial velocity of both stars as measured by an observer far to the right of the figure. At phase ϕ_0 both stars move perpendicularly to the line of sight, at slightly later phases the more massive star moves away from the observer ($v_r > 0$) and the less massive star towards the observed ($v_r < 0$). For observers in a different direction, the radial velocity curves looks the same, but with a phase shift, i.e. a different value for ϕ_0 .

By equating this age with the age of the solar system, we make three assumptions. a) the planetoids solidified soon after the birth of the solar system b) the planetoids all had the same ratio $^{87}\text{Sr}(0)/^{86}\text{Sr}$. c) the meteorites did not melt again. Assumptions b) and c) can be checked directly from the fit: if they are not correct, the meteorites will not lie on the same line.

3.5.2 The radial velocity of a binary star

Consider two masses, M and m , which form a binary. In general, each star moves in an ellipse around the center of gravity, the more massive star in a small ellipse and the less massive star in a large ellipse. We consider the special case of a circular orbit, i.e. the distance a between the stars is constant, and the distance of each star to the center of gravity is also constant, and inversely proportional to the mass of the star. According to the laws of Newton

$$\left(\frac{2\pi}{P}\right)^2 = \frac{G(M+m)}{a^3} \quad (91)$$

where P is the orbital period and

$$a_* = \frac{m}{M+m}a; \quad a_p = \frac{M}{M+m}a \quad (92)$$

The orbits for a binary with $M = 2m$ are shown in Figure 11.

The velocity of the star in its orbit is

$$v_* = \frac{2\pi a_*}{P} = \frac{2\pi a}{P} \frac{m}{M+m} = \frac{m}{M+m} \left[\frac{2\pi G(M+m)}{P} \right]^{1/3} \quad (93)$$

For a far-away observer in the plane of the orbit, the radial velocity of the star varies sinusoidally. For an observer outside the plane, the radial velocities are diminished

by the projection factor $\sin i$, where i is the angle between the normal on the plane and the direction towards the observer: the observer in the plane measures the full orbital velocity, the observer perpendicular to the plane measure zero orbital velocity. The radial velocity of the star as measured by a far-away observer follows:

$$v_{*r} = v_\gamma + v_* \sin(\omega t + \phi_o) \sin i \quad (94)$$

where v_γ is the radial velocity of the center of gravity of the binary, and $\omega \equiv 2\pi/P$.

Computer Exercise 5.

a. you obtain a table with the ratios $^{87}\text{Sr}(t)/^{86}\text{Sr}$ and $^{87}\text{Rb}(t)/^{86}\text{Sr}$, and the errors in these ratios, for 26 meteorites. The errors in the first ratio are smaller (both absolutely and relatively) than those in the second ratio. Argue that use of a standard χ^2 -fit of a line in the form $y = a + bx$ demands that one takes the first ratio as x and the second ratio as y . Rewrite Eq. 90 accordingly.

b. write a computer code that reads the numbers in the table into arrays, and that plots the numbers in a figure, with error bars.

c. find a Python routine to fit a straight line with the least-squares method, and read its description carefully. Use this routine to determine the best fit to all 26 points. Let the code plot this fit in the figure, and in a separate frame also plot for all points the difference between the observed value and fitted values in terms of $(\text{O-C})/\sigma = \chi$ for all points. Plot error bars for the χ 's.

d. Compute the probability that the fit is a good model to the data. Which points have deviations larger than expected by chance? Give a possible reason for the deviation of these points. Remove the four most deviating points (these are the last 4 in the data-file), and fit a straight line to the remaining 22 points. Plot the data with fit, and a frame with the χ 's. Is the probability acceptable? Determine the age of the Earth and its error according to this fit.

Computer Exercise 6.

a. you obtain a table with times (in days), velocities (in m/s) and velocity errors (in m/s) of velocity measurements of 51 Peg, the first star around which a planet was discovered. Make a plot showing these velocities.

b. Write the velocity of 51 Peg as

$$v(t) = a_1 + a_2 \sin(a_3 t + a_4) \quad (95)$$

From the plot, estimate the values of the parameters a_{1-4} . Think of the units!

c. Begin the transition to the Levenberg-Marquardt fitting procedure by formally writing $v(t)$ as $y_m(x, \vec{a})$. The procedure needs a subroutine, which you can call for example `velo`, which for given parameters \vec{a} and argument t computes the velocity y_m and also the four derivatives $\partial y_m / \partial a_i$. The first line of the subroutine should have the form

```
subroutine velo(t,a,y,dyda,na)
```

where `t` and `a(1:na)` are input time and parameter array, with `na` the number of parameters, and output `y` is the velocity at time t , and `dyda(1:na)` the derivatives of the velocity with respect to the parameters. Test your routine by plotting the velocity curve with the parameters from b) in the figure made in a).

d. Read carefully the description of subroutine `mrqmin` in Chapter 15.5 of *Numerical recipes*, and the example function `fgauss`. Note how `fgauss` is written to minimize

repeat calculations, and adapt your routine `velo` in a similar way.

e. Complete your programme so that it fits the velocities of 51 Peg, and writes to the screen the parameter values. First write it with one call to `mrqmin`, and verify that the χ^2 does become smaller. Then write the code so that after each call to `mrqmin` you can choose whether to do another call or not.

f. Finalize the code so that it plots the final solution with the data, and writes to the screen i) the best parameter values, ii) the errors in these values and iii) the probability that your fit is a good description of the data. Also compute the best period and its error

g. Compute the mass of the planet and its error. For this you may assume that the 51 Peg has the same mass as the Sun: $M = M_{\odot} \simeq 2 \times 10^{30}$ kg, and that the planet has a much smaller mass: $m \ll M$.

3.6 What is wrong?

Figure 12 shows a description of a fitting procedure applied to an X-ray spectrum: a truly remarkable piece of bungling. If we have N measurements and M parameters, this is equivalent to N equations for M unknowns. If the measurements are independent, we can solve the parameters uniquely for $N = M$. If $N > M$, there are many solutions, and we pick the one with the maximum likelihood. If $N < M$, there is not enough information to determine the values of M . The ‘solution’ for this problem chosen by the authors is nonsense mathematically, and thus physically.

Note that these conclusions can be drawn without any knowledge of the astrophysics involved.

Since we could not model the spectrum successfully with a two-component model, we introduced an additional component to account for the excess at the very low energy channels. This model consists of a power-law, a soft X-ray excess, and an ultrasoft component. As such a three-component fit requires more parameters than can be justified by the spectral resolution of the PSPC, the fits were performed iteratively where only two parameters were allowed to vary during each step. After these free parameters were optimized, they were fixed in the next fitting round in which two other parameters were treated as free parameters. The adopted ultrasoft com-

Figure 12: *Description of fitting procedure from ApJ 402, 514 (1993)*

4 Errors distributed according to a Poisson distribution and Maximum likelihood

In measurements in which a small number of events is counted, we are often dealing with a distribution of the measurements which is given by the Poisson distribution. In this case, the least-squares methods do not apply. We will discuss the maximum-likelihood method for errors with a Poissonian distribution. In the literature, this method is often referred to as *the maximum-likelihood method*, where the Poissonian error distribution is implied. We will follow this practice, but only after noting that this abbreviated name is misleading, in that the least-squares method is also a maximum-likelihood method, viz. the maximum-likelihood method for errors distributed as Gaussians!

To illustrate the maximum-likelihood method we consider the number of counts on a photon-counting detector. We write the number of counts detected in location i as n_i , and the number of counts predicted for that location by the model as m_i . The probability at one location to obtain n_i photons when m_i are predicted is

$$P_i = \frac{m_i^{n_i} e^{-m_i}}{n_i!} \quad (96)$$

If the values of m_i (and n_i) are large, this probability may be approximated with a Gaussian, and one may use a least-squares method. This is based on the assumption that the difference between the Poisson and Gaussian distributions for large μ is less important than the uncertainties due to systematic effects in the measurements. In principle, this assumption should be verified in each case. Typically, a value of 20 is used; see Fig. 7.

For small values of m_i (and n_i) we proceed to use the Poisson distribution. To maximize the overall probability we thus have to maximize

$$L' \equiv \prod_i P_i \quad (97)$$

and for computational ease we maximize the logarithm of this quantity

$$\ln L' \equiv \sum_i \ln P_i = \sum_i n_i \ln m_i - \sum_i m_i - \sum_i \ln n_i! \quad (98)$$

The last term in this equation doesn't depend on the assumed model, and thus – in terms of selecting the best model – may be considered as a constant. Thus maximizing L' is equivalent to minimizing \mathcal{L} , where

$$\mathcal{L} \equiv -2 \left(\sum_i n_i \ln m_i - \sum_i m_i \right) \quad (99)$$

If one compares two models A and B, with number of fitted parameters M_A and $M_B > M_A$ and with values of \mathcal{L}_A and \mathcal{L}_B , respectively, the difference $\mathcal{L}_B - \mathcal{L}_A$ is a χ^2 distribution with $M_B - M_A$ degrees of freedom, for a sufficient number of photons (Cash 1979 ApJ 228, 939; Mattox et al. 1996 ApJ 461, 396). Table 2 in Section 3.4 can be used for this.

4.1 Some analytical examples

4.1.1 A constant background

As an example we model a constant background: $m_i = A$. With N pixels we thus have a total number of photons in the model $N_m = NA$; the total observed number of photons is $N_o = \sum_{i=1}^N n_i$. Inserting this in eq. 99 we have

$$-0.5 \ln L = \sum n_i \ln A - \sum A = N_o \ln A - NA \quad (100)$$

and we have to minimize this with respect to A . Thus

$$\frac{\partial \ln L}{\partial A} = 0 \Rightarrow \frac{N_o}{A} - N = 0 \Rightarrow N_o = NA = N_m \quad (101)$$

and we see that the best solution, as expected, is the one in which the number of photons in the model equals that in the observation.

4.1.2 Constant background plus one source

Let us now assume that we have a source with strength B , and that a fraction f_i of this lands on pixel i . Eq. 99 now becomes

$$-0.5 \ln L = \sum_i n_i \ln(A + Bf_i) - \sum_i (A + Bf_i) \quad (102)$$

We search the minimum of L for variations in A and B :

$$\frac{\partial \ln L}{\partial A} = 0 \Rightarrow \sum_i \frac{n_i}{A + Bf_i} - \sum_i (1) = \sum_i \frac{n_i}{A + Bf_i} - N = 0 \quad (103)$$

$$\frac{\partial \ln L}{\partial B} = 0 \Rightarrow \sum_i \frac{n_i f_i}{A + Bf_i} - \sum_i f_i = \sum_i \frac{n_i f_i}{A + Bf_i} - 1 = 0 \quad (104)$$

We thus have two equations for the two unknowns A and B . Multiply the first equation with A , the second with B , and add the two equations to find

$$\sum_i n_i = AN + B \quad (105)$$

i.e. the total number of counts in the best model is equal to the total number of observed counts. This condition may be used so that one only has to fit one, rather than two parameters.

4.2 Optimal extraction of a spectrum

For this topic we refer to the article by K.D. Horne, 1985, An optimal extraction algorithm for CCD spectroscopy, PASP 98, 609-617.

4.2.1 Average of two measurements

One of the interesting details in the article by Horne is that after a first iteration, once a model has been computed, the error in the number of counts is computed from the model, and not from the data. To illustrate why, we return briefly to the topic of Sect. 1.5.1. Consider two measurements N_1 and N_2 . *Assuming that both are drawn from the same Gaussian distribution* with average m and width σ , we want to compute the most probable value for m . From Eq. 1 we obtain the combined probability as

$$L = \frac{1}{\sigma\sqrt{2\pi}} e^{-(N_1-m)^2/2\sigma^2} \frac{1}{\sigma\sqrt{2\pi}} e^{-(N_2-m)^2/2\sigma^2} \quad (106)$$

To find the most probable m we set the derivative of L with respect to m equal to zero, and for computational simplicity we first take the logarithm:

$$\frac{d \ln L}{dm} = 0 = \frac{N_1 - m}{\sigma^2} + \frac{N_2 - m}{\sigma^2} \quad \Rightarrow \quad m = \frac{N_1 + N_2}{2} \quad (107)$$

The answer is the same as that for Poissonian distributions (Sect. 1.5.1).

Now suppose we would have estimated the error in each measurement with its square root: $\sigma_1 = \sqrt{N_1}$ and $\sigma_2 = \sqrt{N_2}$, then the smaller measurement would be assigned a smaller error, and an average computed with Eq. 63 would lie closer to the smallest value. This shows how important it is to be clear about the assumptions made before doing a statistical calculation. If we *assume* that both measurements were drawn from the same (Gaussian or Poissonian) distribution, this implies that both have the same weight in determining the average. If we *allow* that the measurements are drawn from distributions with different widths, and *assume* that each width is best estimated from the corresponding measurement, then the measurements with a smaller value will have a higher weight.

4.3 What is wrong?

Figure 13 shows fits to the distributions of the visual luminosity (expressed in absolute magnitude M_V) of clusters in the galaxies M 51 and NGC 6964. The authors clearly have no idea what they are doing. *First*: the reduced χ^2 in each frame is much smaller than 1. This indicates that something is wrong. *Second*: if we look at the graphs, we see that many bins have numbers less than 10. This implies that χ^2 statistics do not apply. *Third*: if we apply Gaussian statistics to a Poisson case, we obtain reduced χ^2 larger than 1: Poisson distributions are much more forgiving for large deviations than Gaussian distributions. Therefore, the authors have not done the χ^2 fit correctly.

Note that these conclusions can be drawn without any knowledge of the astrophysics involved.

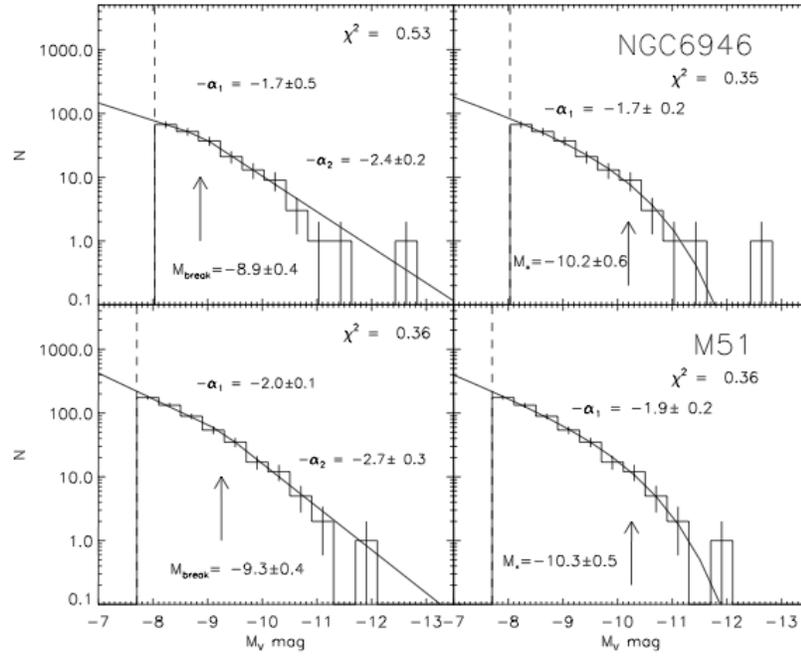


Figure 13: *Four fits to the number of clusters as a function of their absolute magnitude, from A&A 450, 129 (2006).*

5 General Methods

To produce the best fit for the case of Poissonian errors, several methods are available, of which we discuss a few. The methods are distinct from the methods discussed above for the least-squares cases, in that they do not use the derivative of the function, but only the function value itself, together with the criterion of best fit. These methods are equally suited for least-squares problems, with Eq. 60 as the criterion, and for maximum-likelihood problems with Poisson statistics, where Eq. 99 must be minimized. For simple problems, these methods tend to be slower in converging on a good solution than the Levenberg-Marquardt method. For problems with many variables, where the matrix α (Eq. 81) becomes very big, and for problems in which the χ^2 distribution has many local minima, the methods discussed below are actually more efficient than the Levenberg-Marquardt method. An example is the fit to the lightcurves (in U, B, V) and radial velocities of a binary: this problem has many variables, and can have many local minima.

5.1 Amoeba

See Numerical Recipes (2nd ed.) chapter 10.4

5.2 Genetic algorithms

For an explanation of the principle of genetic algorithms we refer to P. Charbonneau, 1995, Genetic algorithms in astronomy and astrophysics, ApJ Supl.Ser. 101, 309-334 (in particular the first 9 pages).

An interesting extension is the use of *black sheep*, i.e. bad descendents from good parents, in the genetic algorithm scheme. This is discussed by A. Bobinger, 2000, Genetic eclipse mapping and the advantage of Black Sheep, A&A 357, 1170-1180

5.3 Computer exercises

We fit the same data as in the last exercise, first with the routine `amoeba` from *Numerical Recipes* and then with the Genetic Algorithm method. We combine two methods for the fits to a two-planets system.

Computer Exercise 7. a. Copy from your code for the previous exercise the part that reads and plots the data into a new program. Now adapt the new program in such a way that the data (number of points, times, velocities, velocity errors) are in a module.

b. Write a function (e.g. `chikwa(a)`) which returns the χ^2 for input parameters `a`. The first line should have the form

```
real function chikwa(a)
```

and the function should obtain the observations via the module from a). The model calls a function which for given time returns the model velocity. Check with the results of the previous exercise that your routine returns correct values for χ^2 .

c. Now read the description in *Numerical Recipes* for the use of `amoeba`, and implement it. Show that it iterates to the same solution and the previous code.

Computer Exercise 8. From the website for this lecture copy `tpikaia.f90`. Have a look at it. The first part is a module to transfer velocity data, the next two parts are a subroutine `velocity` which provides a velocity for input parameters \vec{a}

and a function `chisq` which computes χ^2 for input parameters \vec{x} . You have written similar routines yourself for earlier exercises. Note that the function `chisq` differs in two ways from those earlier versions. First, the parameters \vec{a} are passed on *normalized* to values between 0 and 1; the normalisation factors $a\vec{m}ax$ are passed via a common block. Second, the length `na` of the vectors \vec{x} and $a, a\vec{m}ax$ is passed through the call. This second change is required because `pikaia` expects this form.

The fourth block is the main programme `tpikaia`. This reads the measurements and puts them in the common block. It also computes the normalisation $a\vec{m}ax$ of the parameters \vec{a} that describe the velocity. These two calculations are the only changes you have to make for other applications! (apart from the commands near the end that write the results).

Then follow lines that initialise the routine that provides random numbers (with `rninit`), and the choice for the standard ('default') values for the parameters for the genetic algorithm. When `ctrl(i)` equals -1 the standard value is chosen for the i th parameter. We make one exception by choosing the number of generations explicitly.

Starting with function `urand` the code has been copied directly from `PIKAIA`, with this difference that the main program `xpikaia` and the example function `twod` have been removed. For `PIKAIA` see: Charbonneau, P. 1995, *ApJS* 101, 309 (on the lecture website: `charbonneau.pdf`) and

`www.hao.ucar.edu/Public/models/pikaia/pikaia.html`

a. Compile the code (with some adaptations if necessary) and show that it reproduces the best parameters of 51 Peg.

b. Eq. 95 gives the velocity for a star with one planet. Write the equation for a star with two planets, both in a circular orbit. The data for this exercise are on the website in `psrplan.dat`. This Table gives the differences ΔT between predicted and measured arrival times of the pulses of PSR 1257+12, and can be fitted for two planets with

$$\Delta T(t) = a_1 + a_2 \sin(a_3 t + a_4) + a_5 \sin(a_6 t + a_7)$$

The amplitudes a_2 and a_5 are directly proportional to the velocities of the neutron star due to the two planets A and B:

$$v_{*A} \sin i_A = \frac{2\pi c a_2}{P_A} = a_2 a_3 c \quad \text{and} \quad v_{*B} \sin i_B = \frac{2\pi c a_5}{P_B} = a_5 a_6 c$$

Here $P_{A,B}$ and $\sin_{A,B}$ are the periods and inclinations of the orbits of planets A,B and c is the velocity of light. The velocity amplitudes $v_{*A,B}$ are related to the masses $m_{A,B}$ of the planets via Eq. 93.

Before using the data we make some small but useful adaptations:

1. convert the time differences to milliseconds by multiplying them with 100
2. convert the times to times with respect to the middle of the measurement span as follows:

$$t(i) = t(i) - t_{\text{mid}} \quad \text{where} \quad t_{\text{mid}} \equiv 0.5[t(N) + t(1)]$$

3. all measurements have the same error, viz. 0.1 millisecond.

The first two adaptations minimize the roundoff errors in the computations; the second adaptations minimizes correlations between the fitted parameters. Now adapt the code from a) so that it fits the parameters for the pulsar time differences.

c. Use the code and fit the time differences of PSR 1257+12. You will find that the quality of the solution, as expressed in χ^2 , increases only slowly with the number of generations. Study the improvement in the resulting χ^2 as a function of the number of generations.

d. Adapt the Levenbergh-Marquardt code from an earlier exercise for the case of 2 planets; use the result of c) as start solution, and determine i) the best values for the parameters that describe the orbit ii) the errors in these parameters and iii) the probability that your fit describes the orbit correctly.

e. Answer the following two questions: e1) what is the use of first working with genetic algorithms and only after that use the Levenbergh-Marquardt method? e2) why is it useful to follow up the determination with the genetic algorithm with the Levenbergh-Marquardt method?

Your answer differs from the literature values because, one, we ignore in this problem the eccentricity of the planet orbits, and two, we ignore the later detected third planet.

6 Random numbers

I have not had time yet to write this section; but a good example of the use of random numbers is given by the genetic algorithm.

6.1 Computer exercises

Computer Exercise 9. In this problem you learn how to use random number generators. The Numerical Recipes library has a random number generator function `ran1(idum)` which returns a random number between 0 and 1 for integer input `idum`. For the first call, `idum` must be negative, e.g. do this by setting `idum=-4321` before calling the function for the first time. The routine itself then keeps track of `idum`: do not touch `idum` yourself after the first call!

On the web site you find two subroutines (and a brief explanation on how to use them): one to bin an array of numbers; and one to prepare the binned array for a plot as a histogram.

a. calculate the theoretical distribution of random points on the Earth as a function of longitude ($0 < l < 2\pi$) and as a function of latitude ($-\pi/2 < b < \pi/2$).

b. now generate 1000 points in random positions on the the Earth, using the numerical recipes function `ran1`, and bin these in 10 bins in l . Compute the number you expect theoretically in each bin, and verify that each bin contains the correct number within the statistical error, by plotting the generated points with the theoretical prediction in one plot.

c. Idem for b . (To simplify matters: approximate the integral over the bin as the product of the width of the bin and the theoretical function value in the center of the bin.)

d. Now repeat b) and c) for 10^5 points and 25 bins, to see how the correspondence improves with the number of generated points.

Computer Exercise 10. This exercise illustrates the – at first sight – weird fact that the best estimated value for a measured quantity may be different from the value given by the measurement! For the quantity we use the parallax of a star: the parallax p is in essence the inverse of the distance d of a star:

$$p(\prime\prime) \equiv \frac{1}{d(\text{pc})}$$

where $\prime\prime$ stands for arcsecond: 1 degree is divided in 3600 arcseconds. The above equation defines the parsec pc; $1 \text{ pc} = 3.086 \cdot 10^{16} \text{ m}$. The typical measurement accuracy of the Hipparcos satellite is about 1 milliarcsecond ($\text{mas} \equiv 0.001\prime\prime$).

a. assume that near the Sun the stars are distributed homogeneously in space. Write a computer code that puts stars randomly in a volume with radius $R = 1000$ around the Sun and select those with distances $390 < r(\text{pc}) < 410$. Is the number found with your routine is in agreement with the expected number?

b. Now write a routine that selects a value from the Gaussian error distribution, with width $0.001\prime\prime$, for the parallax of each star. Simulate the parallax measurements of 10 000 stars within a homogeneous volume with radius $R = 1 \text{ kpc}$, and select those with a derived distance between 390 and 410 pc. Compute the real distribution of distances for stars measured in this range, in ten bins with width 100 pc, ranging from 0. to 1000. pc.

c. Compute the expected number of stars in the selected range; does it agree with the number produced with your random routines?

Computer Exercise 11. The numerical recipes function `gasdev(idum)` returns a random number chosen from a Gaussian distribution. Here `idum` has the same function as in the function `ran1`. To illustrate the use of this we produce a Maxwellian three-dimensional velocity distribution by combining three one-dimensional Gaussian velocity distributions, in the x , y and z directions.

a. use the `gasdev` routine to generate 1000 velocities chosen from a Gaussian distribution; bin these in ten bins, ranging from -3σ to $+3\sigma$ and compare with the theoretical distribution in a plot.

b. when you use the binning routine, you will (probably) get some messages on the screen telling you which points lie outside the range within which you binned. With which function can you compute how many such points you expect theoretically? Do so! Is it compatible with what you find?

c. produce 1000 particles, where for each particle you choose a velocity in the x direction, in the y direction and in the z direction, each from a Gaussian distribution, and compute the total velocity. Bin the total velocities in 10 bins, and show that they follow the predicted Maxwell distribution.